ABSTRACT

The invention includes a computer implemented process to identify at least one pattern and its distribution in a set of data for the purpose of interpreting the data. The process comprises (a) representing a set of data by an original data matrix D residing in a storage device, and; (b) decomposing the set of data into a set of patterns represented by a matrix F and their distribution represented by a matrix A, wherein the matrix F represents the set of patterns needed to describe the data and the matrix A represents the distribution of the set of patterns within the data matrix D, the decomposing comprising performing a Bayesian-based Monte Carlo calculation using at least the data matrix D to determine the matrices A and F, wherein the matrices A and F reconstruct the data matrix D and are more amenable to analysis than the data matrix D. Application of the process to environmental, biological and medical, econometric, and other fields is included in the invention.

5

10

15

20

25

30   38662v1